



Bachelor Level / First Year/ First Semester/ Science
Computer Science and Information Technology Stat. 108.
(Statistics I)

Full Marks: 60
Pass Marks: 24
Time: 3 hours.

Candidates are required to give their answers in their own words as far as practicable.
All notations have the usual meanings.

Group A

(2×10=20)

Attempt any Two:

1. Describe simple random sampling with and without replacement for drawing a random sample of size n from a population of size N . In both cases show that sample mean is unbiased estimate of the population mean. Derive the variance of the sample mean in both cases. If V_{with} and V_{without} correspondingly denote that variance of sample mean under simple random sampling with and without replacement method, then show that

$$(V_{\text{with}} - V_{\text{without}}) = \frac{n-1}{Nn} S^2$$

and write conclusion that you can draw out of the above result.

2. Describe function and method of sign test. A study was designed to determine the effect of a certain movie on the moral attitude of young children. The data below represent a rating from 0 to 20 on a moral attitude scale recorded before and after viewing the movie, where high score associated to high morality. Carry out the test of hypothesis that movie had no effect on moral attitude of children against it had using sign test at level 0.1.

Before	14	16	15	18	15	17	19	17	17	16	14	15
After	13	18	16	17	16	19	20	18	19	15	18	16

3. A sample of $n = 22$ data points was used to estimate β_0 , β_1 and β_2 of the multiple regression model: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$, where Y = sales of the product in thousand Rs, X_1 = radio advertising expenditure in thousand Rs and X_2 = newspaper advertising expenditure in thousand Rs. The available estimates of parameters and their standard errors are summarized below:

	β_0	β_1	β_2
Estimate	156.43	13.08	16.80
Standard error	126.76	1.76	2.96

- (a) Write the estimated regression model and interpret the coefficients of X_1 and X_2
(b) Predict the value of Y when $X_1 = 75$ and $X_2 = 50$
(c) Further computation shows that $\sum (Y_i - \bar{Y})^2 = 2507793$ and $SSR = 20280$
 $\sum (Y_i - \hat{Y}_i)^2 = 479760$. SSR
Based on this result compute R^2 and carry out overall significance test of the model
(d) Carry out the test of $H_0: \beta_1 = 0$ against $H_1: \beta_1 > 0$.

Group B

Answer any eight questions:

(8×5=40)

4. In a stratified random sampling, if the cost of survey is constant for each stratum then derive an expression for n_h under optimum allocation.
5. Describe in detail two-stage sampling method. Obtain an expression for an unbiased estimator of population total when samples were drawn by adopting simple random sampling without replacement method, and what would be the expression of this unbiased estimator if $M_i = M$ and $m_i = m$ for all i ?
6. Define a run. A true-false examination constructed with the answers running in the following sequence

TEFTFTFTTFTFTFTFTTE

Does this sequence indicate a departure from randomness in the arrangement of T and F answers? Critical region at 5% level of significance is $R \leq 6$ or $R \geq 16$ where R = number of runs.

7. A survey of voter sentiment was conducted in four wards to compare the fraction of voters candidate A. Random samples of two hundred voters were polled in each of the four wards with result as shown below

	Ward			
	1	2	3	4
Favor A	76	53	59	48
Do not Favor A	124	147	141	152

Do the data present sufficient evidence to indicate that the fraction of voters favoring candidate A differ in the four wards? Use chi-square test at 5% level.

8. The final exam scores obtained by two groups of students, where students of group A were taught using method A and those of group B were taught using method B, are summarized below. Use Mann Whitney test to determine whether or not the final exam scores of two groups are different.

Group A	55	59	61	64	64	70	73	75	76	82	83	95
Group B	65	77	80	80	84	86	88	91	91	91		

9. Describe rationale and method of Median test test.
10. Describe the method of estimation of the parameters α and β of the growth model $Y_t = \alpha e^{\beta t}$, when the values of Y_t are available for $t = 1, 2, \dots, n$? What does the estimated β measure?
11. Describe the method of formulating a multiple regression model when the dependent variable Y is binary in nature.
12. Write a suitable multiple regression model if the dependent variable (Y) is output and one independent variable (X_1) is labor input and other independent variable (X_2) is capital input. Describe the properties of the model.
13. Define multiple correlation coefficient, If $r_{12} = 0.679$, $r_{13} = 0.502$ and $r_{23} = -0.092$, then compute $R_{1.23}$, where r_{ij} is the simple correlation coefficient between X_1 , and X_2 and X_3 .
14. What is autocorrelation? How do you estimate and test it?